

# A SIMPLE PROOF OF STOLARSKY'S INVARIANCE PRINCIPLE\*

Johann S. Brauchart<sup>†</sup> and Josef Dick<sup>‡</sup>  
 School of Mathematics and Statistics,  
 University of New South Wales,  
 Sydney, NSW, 2052, Australia  
 j.brauchart@unsw.edu.au,  
 josef.dick@unsw.edu.au

## Abstract

Stolarsky [Proc. Amer. Math. Soc. 41 (1973), 575–582] showed a beautiful relation that balances the sums of distances of points on the unit sphere and their spherical cap  $\mathbb{L}_2$ -discrepancy to give the distance integral of the uniform measure on the sphere a potential-theoretical quantity (Björck [Ark. Mat. 3 (1956), 255–269]). Read differently it expresses the worst-case numerical integration error for functions from the unit ball in a certain Hilbert space setting in terms of the  $\mathbb{L}_2$ -discrepancy and vice versa (first author and Womersley [Preprint]). In this note we give a simple proof of the invariance principle using reproducing kernel Hilbert spaces.

## 1 Introduction

We consider the unit sphere

$$\mathbb{S}^d = \left\{ \mathbf{z} = (z_1, \dots, z_{d+1}) \in \mathbb{R}^{d+1} : \|\mathbf{z}\| = \sqrt{z_1^2 + \dots + z_{d+1}^2} = 1 \right\}$$

embedded in the Euclidean space  $\mathbb{R}^{d+1}$ ,  $d \geq 2$ . Let  $f : \mathbb{S}^d \rightarrow \mathbb{C}$  be a continuous function. Then we approximate the integral  $\int_{\mathbb{S}^d} f(\mathbf{x}) \, d\sigma_d(\mathbf{x})$ , where  $\sigma_d$  is the normalized Lebesgue surface area measure on  $\mathbb{S}^d$  ( $\int_{\mathbb{S}^d} d\sigma_d = 1$ ), by an equal weight numerical integration rule

$$Q_N(f) = \frac{1}{N} \sum_{k=0}^{N-1} f(\mathbf{z}_k) \quad (1)$$

where  $\mathbf{z}_0, \dots, \mathbf{z}_{N-1} \in \mathbb{S}^d$  are the integration nodes on the sphere. In order to analyze the integration error committed by the approximation, we define a worst-case error by

$$e(\mathcal{H}, Q_N) = \sup_{f \in \mathcal{H}, \|f\| \leq 1} \left| \int_{\mathbb{S}^d} f(\mathbf{x}) \, d\sigma_d(\mathbf{x}) - \frac{1}{N} \sum_{k=0}^{N-1} f(\mathbf{z}_k) \right|,$$

where  $\mathcal{H}$  denotes a normed function space with norm  $\|\cdot\|$ . The rate of decay of the worst-case error depends on the function space and the integration nodes. For a fixed function space, the

\*MSC: primary 41A30; secondary 11K38, 41A55.

<sup>†</sup>The author is supported by an APART-Fellowship of the Austrian Academy of Sciences.

<sup>‡</sup>The author is supported by an Australian Research Council Queen Elizabeth 2 Fellowship.

worst-case error can serve as a quality criterion for different sets of integration nodes, meaning that the performance of a set of  $N$  quadrature points  $\mathbf{z}_0, \dots, \mathbf{z}_{N-1}$  can be compared to another set of  $N$  quadrature points by comparing the corresponding worst-case errors. Generally, this only means that the integration error is smaller, but sometimes the worst-case error allows also a geometrical interpretation.

Another quality criterion for points on the sphere exploits the potential energy, or more generally, the Riesz  $s$ -energy of configurations of points modeling unit charges which are thought to interact through a potential  $1/\|\cdot\|^s$  ( $s \neq 0$ ), where  $\|\cdot\|$  denotes the Euclidean distance. (We refer the reader to the survey papers [10] and [13] and for *universally optimal configurations* to [8].) A particular instance is the (normalized) sum of distances ( $s = -1$ )

$$\frac{1}{N^2} \sum_{k,\ell=0}^{N-1} \|\mathbf{z}_k - \mathbf{z}_\ell\|, \quad \mathbf{z}_1, \dots, \mathbf{z}_N \in \mathbb{S}^d.$$

It is well-known from potential theory (see Björck [4]) that this (discrete) sum of distances of optimal  $N$ -point configurations approaches the associated (continuous) distance integral of the uniform measure  $\sigma_d$  on  $\mathbb{S}^d$  as  $N \rightarrow \infty$ . In fact, any sequence of  $N$ -point systems with this property turns out to be 'asymptotically uniformly distributed'; that is, the discrete probability measure obtained by placing equal charges at the points tends to the uniform measure in the weak-star sense. The difference

$$\int_{\mathbb{S}^d} \int_{\mathbb{S}^d} \|\mathbf{z} - \mathbf{x}\| d\sigma_d(\mathbf{z}) d\sigma_d(\mathbf{x}) - \frac{1}{N^2} \sum_{k,\ell=0}^{N-1} \|\mathbf{z}_k - \mathbf{z}_\ell\| \quad (2)$$

measuring the deviation between theoretical and empirical  $(-1)$ -energy quantifies the quality of points on the sphere (and, indirectly, their uniform distribution) using energy. It should be mentioned that the upper bound of correct order  $N^{-1-1/d}$  for (2) (for optimal configurations) was obtained by Stolarsky [16] using his invariance principle and a result of Schmidt [14] on the discrepancy of spherical caps. The correct-order lower bound ( $N^{-1-1/d}$ ) was established by Beck [3] using his Fourier transform technique.

The spherical cap discrepancy measures the maximum deviation between theoretical and empirical distribution with respect to spherical caps as test sets. It can be used to compare point sets on the sphere with respect to their distribution properties. To introduce the concept of spherical cap discrepancy, we require some notation. A spherical cap centered at  $\mathbf{x} \in \mathbb{S}^d$  with 'height'  $t \in [-1, 1]$  is the set

$$C(\mathbf{x}; t) = \{\mathbf{z} \in \mathbb{S}^d : \langle \mathbf{x}, \mathbf{z} \rangle \geq t\}.$$

The family of all spherical caps is denoted by

$$\mathcal{C} = \{C(\mathbf{x}; t) : \mathbf{x} \in \mathbb{S}^d, t \in [-1, 1]\}.$$

For a set  $J \subseteq \mathbb{R}^{d+1}$  we define the indicator function

$$1_J(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in J, \\ 0 & \text{otherwise.} \end{cases}$$

For a measurable set  $J \subseteq \mathbb{S}^d$  let

$$\sigma_d(J) = \int_{\mathbb{S}^d} 1_J(\mathbf{x}) d\sigma_d(\mathbf{x}).$$

Using spherical caps we can define another quality criterion for points on the sphere  $\mathbb{S}^d$  in terms of their distribution properties. One such criterion is the spherical cap  $L_2$ -discrepancy, which is given by

$$L_2(P) = \left( \int_{-1}^1 \int_{\mathbb{S}^d} \left| \sigma_d(C(\mathbf{z}; t)) - \frac{1}{N} \sum_{k=0}^{N-1} 1_{C(\mathbf{z}; t)}(\mathbf{z}_k) \right|^2 d\sigma_d(\mathbf{z}) dt \right)^{1/2},$$

where  $P = \{\mathbf{z}_0, \dots, \mathbf{z}_{N-1}\}$ .

We considered three, seemingly different, measures which can be applied to point sets on the sphere. It turns out that in some instances, the three measures are related to each other. Stolarksy's insight [16] was that the sum of distances of points on the sphere and the spherical cap discrepancy coincide. On the other hand, also the sum of distances of points on the sphere and the worst-case error coincide for a certain choice of function space, see [7] and also Sloan and Womersley [15] regarding a generalized discrepancy of Cui and Freeden [9].

In this paper we give a simple proof of these results based on reproducing kernel Hilbert spaces. We also provide some generalizations which follow from our approach.

## 2 Reproducing kernel Hilbert space

We define a reproducing kernel Hilbert space using the general approach of [12, Ch. 9.6].

For  $\mathbf{x}, \mathbf{y} \in \mathbb{S}^d$  we define the function  $K_{\mathcal{C}} : \mathbb{S}^d \times \mathbb{S}^d \rightarrow \mathbb{R}$  by

$$K_{\mathcal{C}}(\mathbf{x}, \mathbf{y}) = \int_{-1}^1 \int_{\mathbb{S}^d} 1_{C(\mathbf{z}; t)}(\mathbf{x}) 1_{C(\mathbf{z}; t)}(\mathbf{y}) d\sigma_d(\mathbf{z}) dt. \quad (3)$$

Since  $1_{C(\mathbf{z}; t)}(\mathbf{x}) = 1_{C(\mathbf{x}; t)}(\mathbf{z})$ , we also have

$$K_{\mathcal{C}}(\mathbf{x}, \mathbf{y}) = \int_{-1}^1 \int_{\mathbb{S}^d} 1_{C(\mathbf{x}; t)}(\mathbf{z}) 1_{C(\mathbf{y}; t)}(\mathbf{z}) d\sigma_d(\mathbf{z}) dt.$$

The function is obviously symmetric, i.e. we have  $K_{\mathcal{C}}(\mathbf{x}, \mathbf{y}) = K_{\mathcal{C}}(\mathbf{y}, \mathbf{x})$ . Further, let  $a_0, \dots, a_{N-1} \in \mathbb{C}$  and  $\mathbf{x}_0, \dots, \mathbf{x}_{N-1} \in \mathbb{S}^d$ . Then we have

$$\begin{aligned} \sum_{k, \ell=0}^{N-1} a_k \overline{a_{\ell}} K_{\mathcal{C}}(\mathbf{x}_k, \mathbf{x}_{\ell}) &= \int_{-1}^1 \int_{\mathbb{S}^d} \sum_{k, \ell=0}^{N-1} a_k \overline{a_{\ell}} 1_{C(\mathbf{x}; t)}(\mathbf{x}_k) 1_{C(\mathbf{x}; t)}(\mathbf{x}_{\ell}) d\sigma_d(\mathbf{x}) dt \\ &= \int_{-1}^1 \int_{\mathbb{S}^d} \left| \sum_{k=0}^{N-1} a_k 1_{C(\mathbf{x}; t)}(\mathbf{x}_k) \right|^2 d\sigma_d(\mathbf{x}) dt \\ &\geq 0. \end{aligned}$$

Thus, the function  $K_{\mathcal{C}}$  is symmetric and positive definite. By [2], this implies that  $K_{\mathcal{C}}$  is a reproducing kernel. It is also shown in [2] that a reproducing kernel uniquely defines a Hilbert space of functions with a certain inner product. Let  $\mathcal{H}_{\mathcal{C}} = \mathcal{H}(K_{\mathcal{C}}, \mathbb{S}^d)$  denote the corresponding reproducing kernel Hilbert space of functions  $f : \mathbb{S}^d \rightarrow \mathbb{R}$  with reproducing kernel  $K_{\mathcal{C}}$ .

We consider now functions  $f_1, f_2 : \mathbb{S}^d \rightarrow \mathbb{C}$  which permit a certain integral representation. Let  $g_1, g_2 : \mathbb{S}^d \times [-1, 1] \rightarrow \mathbb{C}$  with  $g_1, g_2 \in \mathbb{L}_2(\mathbb{S}^d \times [-1, 1])$  and

$$f_i(\mathbf{x}) = \int_{-1}^1 \int_{\mathbb{S}^d} g_i(\mathbf{z}; t) 1_{C(\mathbf{z}; t)}(\mathbf{x}) d\sigma_d(\mathbf{z}) dt, \quad i = 1, 2. \quad (4)$$

Notice that for any fixed  $\mathbf{y} \in \mathbb{S}^d$  the function  $K_{\mathcal{C}}(\cdot, \mathbf{y})$  also is of this form, where the function  $g$  is given by  $1_{C(\mathbf{z};t)}(\mathbf{y})$  (considered as a function of  $\mathbf{z}$  and  $t$  and where  $\mathbf{y}$  is fixed). For functions of this form we can define an inner product by

$$\langle f_1, f_2 \rangle_{K_{\mathcal{C}}} = \int_{-1}^1 \int_{\mathbb{S}^d} g_1(\mathbf{z}; t) \overline{g_2(\mathbf{z}; t)} d\sigma_d(\mathbf{z}) dt. \quad (5)$$

Let  $\mathbf{y} \in \mathbb{S}^d$  be fixed. With this definition we obtain

$$\langle f_1, K_{\mathcal{C}}(\cdot, \mathbf{y}) \rangle_{K_{\mathcal{C}}} = \int_{-1}^1 \int_{\mathbb{S}^d} g_1(\mathbf{z}; t) 1_{C(\mathbf{z};t)}(\mathbf{y}) d\sigma_d(\mathbf{z}) dt = f_1(\mathbf{y}).$$

By [2], the inner product in  $\mathcal{H}_{\mathcal{C}}$  is unique. Therefore, functions  $f_i$ , which are given by (4) and for which  $\langle f_i, f_i \rangle_{K_{\mathcal{C}}} < \infty$ , are in  $\mathcal{H}_{\mathcal{C}}$  and (5) is an inner product for those functions in  $\mathcal{H}_{\mathcal{C}}$ .

Consider now the reproducing kernel  $K_{\mathcal{C}}$ . We have

$$\int_{-1}^1 1_{C(\mathbf{z};t)}(\mathbf{x}) 1_{C(\mathbf{z};t)}(\mathbf{y}) dt = \int_{-1}^{\min\{\langle \mathbf{x}, \mathbf{z} \rangle, \langle \mathbf{y}, \mathbf{z} \rangle\}} dt = 1 + \min\{\langle \mathbf{x}, \mathbf{z} \rangle, \langle \mathbf{y}, \mathbf{z} \rangle\}.$$

Thus

$$\begin{aligned} K_{\mathcal{C}}(\mathbf{x}, \mathbf{y}) &= \int_{-1}^1 \int_{\mathbb{S}^d} 1_{C(\mathbf{z};t)}(\mathbf{x}) 1_{C(\mathbf{z};t)}(\mathbf{y}) d\sigma_d(\mathbf{z}) dt \\ &= 1 + \int_{\mathbb{S}^d} \min\{\langle \mathbf{x}, \mathbf{z} \rangle, \langle \mathbf{y}, \mathbf{z} \rangle\} d\sigma_d(\mathbf{z}). \end{aligned}$$

We have

$$\min\{\langle \mathbf{x}, \mathbf{z} \rangle, \langle \mathbf{y}, \mathbf{z} \rangle\} = \frac{1}{2} [|\langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle - |\langle \mathbf{x} - \mathbf{y}, \mathbf{z} \rangle|]$$

and

$$\int_{\mathbb{S}^d} \langle \mathbf{x}, \mathbf{z} \rangle d\sigma_d(\mathbf{z}) = 0.$$

If  $\mathbf{x} = \mathbf{y}$ , then we therefore obtain

$$\int_{\mathbb{S}^d} \min\{\langle \mathbf{x}, \mathbf{z} \rangle, \langle \mathbf{y}, \mathbf{z} \rangle\} d\sigma_d(\mathbf{z}) = 0.$$

Let now  $\mathbf{x} \neq \mathbf{y}$ . Then

$$\begin{aligned} \int_{\mathbb{S}^d} \min\{\langle \mathbf{x}, \mathbf{z} \rangle, \langle \mathbf{y}, \mathbf{z} \rangle\} d\sigma_d(\mathbf{z}) &= -\frac{1}{2} \int_{\mathbb{S}^d} |\langle \mathbf{x} - \mathbf{y}, \mathbf{z} \rangle| d\sigma_d(\mathbf{z}) \\ &= -\|\mathbf{x} - \mathbf{y}\| \frac{1}{2} \int_{\mathbb{S}^d} \left\langle \frac{\mathbf{x} - \mathbf{y}}{\|\mathbf{x} - \mathbf{y}\|}, \mathbf{z} \right\rangle d\sigma_d(\mathbf{z}). \end{aligned}$$

The last integral does not depend on the unit vector  $(\mathbf{x} - \mathbf{y})/\|\mathbf{x} - \mathbf{y}\|$  by rotational symmetry. Thus, we have (cf. Appendix A)

$$C_d := \frac{1}{2} \int_{\mathbb{S}^d} |\langle \mathbf{p}, \mathbf{z} \rangle| d\sigma_d(\mathbf{z}) = \frac{1}{d} \frac{\omega_{d-1}}{\omega_d} = \frac{\mathcal{H}_d(\mathbb{B}^d)}{\mathcal{H}_d(\mathbb{S}^d)} = \frac{1}{d} \frac{\Gamma((d+1)/2)}{\sqrt{\pi} \Gamma(d/2)} \sim \frac{1}{\sqrt{2\pi d}} \quad \text{as } d \rightarrow \infty. \quad (6)$$

(Here,  $\omega_d$  is the surface area of  $\mathbb{S}^d$ ,  $\mathbb{B}^d$  is the unit ball in  $\mathbb{R}^d$ ,  $\mathcal{H}_d$  is the  $d$ -dimensional Hausdorff measure normalized such that the  $d$ -dimensional unit cube  $[0, 1]^d$  has measure one and  $\Gamma(z)$  is the Gamma function.)

Therefore we obtain the following closed form representation:

$$K_{\mathcal{C}}(\mathbf{x}, \mathbf{y}) = 1 - C_d \|\mathbf{x} - \mathbf{y}\|. \quad (7)$$

The reproducing kernel has the following properties: for  $\mathbf{y}, \mathbf{z} \in \mathbb{S}^d$  we have

- $1 - 2C_d \leq K(\mathbf{x}, \mathbf{y}) \leq 1$ ;
- $K(\mathbf{x}, \mathbf{y}) = 1 - 2C_d$  if and only if  $\mathbf{x} = -\mathbf{y}$  (clearly,  $0 < 2C_d < 1$ );
- $K(\mathbf{x}, \mathbf{y}) = 1$  if and only if  $\mathbf{x} = \mathbf{y}$ ;

Note that the Karhunen-Loevey expansion of the function  $\|\mathbf{x} - \mathbf{y}\|$  is based on ultraspherical harmonics. Hence the eigenfunctions of  $K_{\mathcal{C}}$  are the ultraspherical harmonics. The corresponding eigenvalues are also known. Therefore, the functions in  $\mathcal{H}_{\mathcal{C}}$  can be expanded using ultraspherical harmonics and the inner product can be written using the coefficients of such an expansion. See [7] for these results.

### 3 Worst-case error

Let  $\|f\|_{K_{\mathcal{C}}} = \sqrt{\langle f, f \rangle_{K_{\mathcal{C}}}}$  denote the norm in  $\mathcal{H}_{\mathcal{C}}$ . Then we define the worst-case error for a quadrature rule  $Q_N$  given in (1) by

$$e(\mathcal{H}_{\mathcal{C}}, Q_N) = \sup \left\{ \left| \int_{\mathbb{S}^d} f(\mathbf{x}) \, d\mathbf{x} - Q_N(f) \right| : f \in \mathcal{H}_{\mathcal{C}}, \|f\|_{K_{\mathcal{C}}} \leq 1 \right\}.$$

Let  $f \in \mathcal{H}_{\mathcal{C}}$ . Then, by the reproducing kernel property  $f(\mathbf{y}) = \langle f, K_{\mathcal{C}}(\cdot, \mathbf{y}) \rangle_{K_{\mathcal{C}}}$  for  $\mathbf{y} \in \mathbb{S}^d$ , and, since the integration functional  $f \mapsto \int_{\mathbb{S}^d} f \, d\sigma_d$  is bounded on  $\mathcal{H}_{\mathcal{C}}$  and has the 'representer'  $\int_{\mathbb{S}^d} K_{\mathcal{C}}(\cdot, \mathbf{z}) \, d\sigma_d(\mathbf{z})$ , one can write

$$\int_{\mathbb{S}^d} f(\mathbf{x}) \, d\sigma_d(\mathbf{x}) - \frac{1}{N} \sum_{k=0}^{N-1} f(\mathbf{x}_k) = \langle f, \mathcal{R}(\mathcal{H}_{\mathcal{C}}, Q_N; \cdot) \rangle_{K_{\mathcal{C}}},$$

where the 'representer' of the error of numerical integration for the rule  $Q_N$  for functions in  $\mathcal{H}_{\mathcal{C}}$  is given by

$$\mathcal{R}(\mathcal{H}_{\mathcal{C}}, Q_N; \mathbf{x}) = \int_{\mathbb{S}^d} K_{\mathcal{C}}(\mathbf{x}, \mathbf{y}) \, d\sigma_d(\mathbf{y}) - \frac{1}{N} \sum_{k=0}^{N-1} K_{\mathcal{C}}(\mathbf{x}, \mathbf{x}_k), \quad \mathbf{x} \in \mathbb{S}^d.$$

The Cauchy-Schwarz inequality yields

$$\left| \int_{\mathbb{S}^d} f(\mathbf{x}) \, d\sigma_d(\mathbf{x}) - \frac{1}{N} \sum_{k=0}^{N-1} f(\mathbf{x}_k) \right| = \left| \langle f, \mathcal{R}(\mathcal{H}_{\mathcal{C}}, Q_N; \cdot) \rangle_{K_{\mathcal{C}}} \right| \leq \|f\|_{K_{\mathcal{C}}} \|\mathcal{R}(\mathcal{H}_{\mathcal{C}}, Q_N; \cdot)\|_{K_{\mathcal{C}}}.$$

In particular, equality is assumed in the last relation when taking  $f$  to be the 'representer'  $\mathcal{R}(\mathcal{H}_{\mathcal{C}}, Q_N; \cdot)$  itself. It follows that

$$e(\mathcal{H}_{\mathcal{C}}, Q_N) = \|\mathcal{R}(\mathcal{H}_{\mathcal{C}}, Q_N; \cdot)\|_{K_{\mathcal{C}}} = \left\| \int_{\mathbb{S}^d} K_{\mathcal{C}}(\cdot, \mathbf{x}) \, d\sigma_d(\mathbf{x}) - \frac{1}{N} \sum_{k=0}^{N-1} K_{\mathcal{C}}(\cdot, \mathbf{x}_k) \right\|_{K_{\mathcal{C}}}. \quad (8)$$

Expanding the square of the worst-case error and substituting the closed form of the reproducing kernel we arrive at the well-known representation

$$\begin{aligned} [e(\mathcal{H}_{\mathcal{C}}, Q_N)]^2 &= \langle \mathcal{R}(\mathcal{H}_{\mathcal{C}}, Q_N; \cdot), \mathcal{R}(\mathcal{H}_{\mathcal{C}}, Q_N; \cdot) \rangle \\ &= \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} K_{\mathcal{C}}(\mathbf{x}, \mathbf{y}) \, d\sigma_d(\mathbf{x}) \, d\sigma_d(\mathbf{y}) - \frac{2}{N} \sum_{k=0}^{N-1} \int_{\mathbb{S}^d} K_{\mathcal{C}}(\mathbf{x}, \mathbf{x}_k) \, d\sigma_d(\mathbf{x}) + \frac{1}{N^2} \sum_{k, \ell=0}^{N-1} K_{\mathcal{C}}(\mathbf{x}_k, \mathbf{x}_\ell) \end{aligned}$$

$$= C_d \left[ \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} \|\mathbf{x} - \mathbf{y}\| \, d\sigma_d(\mathbf{x}) \, d\sigma_d(\mathbf{y}) - \frac{1}{N^2} \sum_{k,\ell=0}^{N-1} \|\mathbf{x}_k - \mathbf{x}_\ell\| \right]. \quad (9)$$

This shows how the square worst-case error in our reproducing kernel Hilbert space is related to (2). In the next section we show how the worst-case error  $e(\mathcal{H}_{\mathcal{C}}, Q_N)$  in our reproducing kernel Hilbert space relates to the  $L_2$  spherical cap discrepancy.

## 4 Spherical cap discrepancy and Stolarsky's invariance principle

Using the integral representation of the reproducing kernel (3) we have

$$\int_{\mathbb{S}^d} K_{\mathcal{C}}(\mathbf{x}, \mathbf{y}) \, d\sigma_d(\mathbf{y}) = \int_{-1}^1 \int_{\mathbb{S}^d} 1_{C(\mathbf{z};t)}(\mathbf{x}) \sigma_d(C(\mathbf{z};t)) \, d\sigma_d(\mathbf{z}) \, dt$$

and

$$\frac{1}{N} \sum_{k=0}^{N-1} K_{\mathcal{C}}(\mathbf{x}, \mathbf{x}_k) = \int_{-1}^1 \int_{\mathbb{S}^d} \frac{1}{N} \sum_{k=0}^{N-1} 1_{C(\mathbf{x};t)}(\mathbf{z}) 1_{C(\mathbf{x}_k;t)}(\mathbf{z}) \, d\sigma_d(\mathbf{z}) \, dt.$$

Thus, the 'representer' of the error of numerical integration is of the form (4); that is

$$\mathcal{R}(\mathcal{H}_{\mathcal{C}}, Q_N; \mathbf{x}) = \int_{-1}^1 \int_{\mathbb{S}^d} 1_{C(\mathbf{z};t)}(\mathbf{x}) \left[ \sigma_d(C(\mathbf{z};t)) - \frac{1}{N} \sum_{k=0}^{N-1} 1_{C(\mathbf{x}_k;t)}(\mathbf{z}) \right] \, d\sigma_d(\mathbf{z}) \, dt.$$

Therefore, using the inner product representation (5) in (8), we obtain

$$e(\mathcal{H}_{\mathcal{C}}, Q_N) = \left( \int_{-1}^1 \int_{\mathbb{S}^d} \left| \sigma_d(C(\mathbf{z};t)) - \frac{1}{N} \sum_{k=0}^{N-1} 1_{C(\mathbf{x}_k;t)}(\mathbf{z}) \right|^2 \, d\sigma_d(\mathbf{z}) \, dt \right)^{1/2} \quad (10)$$

stating that the worst-case error of the numerical integration formula  $Q_N$  in (1) in the considered Sobolev space setting equals the so-called spherical cap  $\mathbb{L}_2$ -discrepancy of the integration nodes.

Combining (9) and (10), we arrive at Stolarsky's invariance principle for the Euclidean distance on spheres.

**Proposition 1 (Stolarsky [16])** *Let  $\mathbf{x}_0, \dots, \mathbf{x}_{N-1} \in \mathbb{S}^d$  be an arbitrary  $N$  point configuration on the sphere  $\mathbb{S}^d$ . Then we have*

$$\begin{aligned} & \frac{1}{N^2} \sum_{k,\ell=0}^{N-1} \|\mathbf{x}_k - \mathbf{x}_\ell\| + \frac{1}{C_d} \int_{-1}^1 \int_{\mathbb{S}^d} \left| \sigma_d(C(\mathbf{z};t)) - \frac{1}{N} \sum_{k=0}^{N-1} 1_{C(\mathbf{x}_k;t)}(\mathbf{z}) \right|^2 \, d\sigma_d(\mathbf{z}) \, dt \\ &= \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} \|\mathbf{x} - \mathbf{y}\| \, d\sigma_d(\mathbf{x}) \, d\sigma_d(\mathbf{y}). \end{aligned} \quad (11)$$

The  $\mathbb{L}_2$ -discrepancy of an  $N$ -point configuration on  $\mathbb{S}^d$  decreases as its sum of distances increases and vice versa. The right-hand side is the distance integral of the uniform measure  $\sigma_d$  on the sphere  $\mathbb{S}^d$  which is the unique extremal measure (also known as the equilibrium measure) maximizing the distance integral

$$\mathcal{I}[\mu] := \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} \|\mathbf{x} - \mathbf{y}\| \, d\mu(\mathbf{x}) \, d\mu(\mathbf{y})$$

over the family of (Borel) probability measures  $\mu$  supported on  $\mathbb{S}^d$ . For the potential theory of the generalized distance integral we refer to Björck [4].

## 5 A weighted reproducing kernel

The above results can be generalized by introducing a weight function. Let  $v : [-1, 1] \rightarrow \mathbb{R}$  satisfy  $v(t) > 0$  for all  $t$  and which has an antiderivative, which we denote by  $V$ . Then we define the reproducing kernel with weight function  $v$  as follows

$$K_{\mathcal{C},v}(\mathbf{x}, \mathbf{y}) = \int_{-1}^1 v(t) \int_{\mathbb{S}^d} 1_{C(\mathbf{z};t)}(\mathbf{x}) 1_{C(\mathbf{z};t)}(\mathbf{y}) \, d\sigma_d(\mathbf{z}) \, dt, \quad \mathbf{x}, \mathbf{y} \in \mathbb{S}^d. \quad (12)$$

For functions represented by integrals

$$f_i(\mathbf{x}) = \int_{-1}^1 \int_{\mathbb{S}^d} g_i(\mathbf{z};t) 1_{C(\mathbf{z};t)}(\mathbf{x}) \, d\sigma_d(\mathbf{z}) \, dt, \quad i = 1, 2,$$

the corresponding inner product is now given by

$$\langle f_1, f_2 \rangle_{K_{\mathcal{C},v}} = \int_{-1}^1 \frac{1}{v(t)} \int_{\mathbb{S}^d} g_1(\mathbf{z};t) \overline{g_2(\mathbf{z};t)} \, d\sigma_d(\mathbf{z}) \, dt.$$

The reproducing kernel can be written as

$$K_{\mathcal{C},v}(\mathbf{x}, \mathbf{y}) = \int_{\mathbb{S}^d} V(\min\{\langle \mathbf{x}, \mathbf{z} \rangle, \langle \mathbf{y}, \mathbf{z} \rangle\}) \, d\sigma_d(\mathbf{z}) - V(-1), \quad \mathbf{x}, \mathbf{y} \in \mathbb{S}^d. \quad (13)$$

For certain weight functions  $v$ , this expression may have a concise form. This reproducing kernel defines a reproducing kernel Hilbert space  $\mathcal{H}_{\mathcal{C},v}$ .

The 'representer' of the error of numerical integration for the rule  $Q_N$  for functions in  $\mathcal{H}_{\mathcal{C},v}$  takes on the form

$$\mathcal{R}(\mathcal{H}_{\mathcal{C},v}, Q_N; \mathbf{x}) = \int_{-1}^1 \int_{\mathbb{S}^d} 1_{C(\mathbf{z};t)}(\mathbf{x}) v(t) \left[ \sigma_d(C(\mathbf{z};t)) - \frac{1}{N} \sum_{k=0}^{N-1} 1_{C(\mathbf{x}_k;t)}(\mathbf{z}) \right] \, d\sigma_d(\mathbf{z}) \, dt.$$

We claim that  $K_{\mathcal{C},v}(\mathbf{x}, \mathbf{y})$  is a function of the inner product  $\langle \mathbf{x}, \mathbf{y} \rangle$ , cf. Appendix B. Using the same approach as before we obtain

$$[e(\mathcal{H}_{\mathcal{C},v}, Q_N)]^2 = \frac{1}{N^2} \sum_{k,\ell=0}^{N-1} K_{\mathcal{C},v}(\mathbf{x}_k, \mathbf{x}_\ell) - \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} K_{\mathcal{C},v}(\mathbf{x}, \mathbf{y}) \, d\sigma_d(\mathbf{x}) \, d\sigma_d(\mathbf{y}). \quad (14)$$

This worst case error can also be expressed in terms of a weighted discrepancy measure:

$$\begin{aligned} e(\mathcal{H}_{\mathcal{C},v}, Q_N) &= \left\| \int_{\mathbb{S}^d} K_{\mathcal{C},v}(\cdot, \mathbf{y}) \, d\sigma_d(\mathbf{y}) - \frac{1}{N} \sum_{k=0}^{N-1} K_{\mathcal{C},v}(\cdot, \mathbf{x}_k) \right\|_{K_{\mathcal{C},v}} = \|\mathcal{R}(\mathcal{H}_{\mathcal{C},v}, Q_N; \cdot)\|_{K_{\mathcal{C},v}} \\ &= \left( \int_{-1}^1 v(t) \int_{\mathbb{S}^d} \left| \sigma_d(C(\mathbf{x};t)) - \frac{1}{N} \sum_{k=0}^{N-1} 1_{C(\mathbf{x};t)}(\mathbf{x}_k) \right|^2 \, d\sigma_d(\mathbf{x}) \, dt \right)^{1/2}. \end{aligned} \quad (15)$$

Using (14) and (15) we obtain the weighted version of Stolarsky invariance principle.

**Theorem 1** *Let  $\mathbf{x}_0, \dots, \mathbf{x}_{N-1} \in \mathbb{S}^d$  be an arbitrary  $N$  point configuration on the sphere  $\mathbb{S}^d$ . Let  $K_{\mathcal{C},v}$  be the weighted reproducing kernel given by (12). Then we have*

$$\begin{aligned} &\frac{1}{N^2} \sum_{k,\ell=0}^{N-1} K_{\mathcal{C},v}(\mathbf{x}_k, \mathbf{x}_\ell) + \int_{-1}^1 v(t) \int_{\mathbb{S}^d} \left| \sigma_d(C(\mathbf{x};t)) - \frac{1}{N} \sum_{k=0}^{N-1} 1_{C(\mathbf{x};t)}(\mathbf{x}_k) \right|^2 \, d\sigma_d(\mathbf{x}) \, dt \\ &= \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} K_{\mathcal{C},v}(\mathbf{x}, \mathbf{y}) \, d\sigma_d(\mathbf{x}) \, d\sigma_d(\mathbf{y}). \end{aligned}$$

The double integral above can be expressed in terms of the weight function, see (20).

In [5], Stolarsky's (general) invariance principle is extended and used to get bounds for the spherical cap discrepancy, see also [6]. Stolarsky [17] also extended his principle to certain metric spaces arising from measures.

Stolarsky [16] introduced the function

$$\rho(\mathbf{x}, \mathbf{y}) = \int_{\mathbb{S}^d} \int_{\min\{\langle \mathbf{x}, \mathbf{y} \rangle, \langle \mathbf{y}, \mathbf{t} \rangle\}}^{\max\{\langle \mathbf{x}, \mathbf{y} \rangle, \langle \mathbf{y}, \mathbf{t} \rangle\}} g(u) \, du \, d\sigma_d(\mathbf{z}), \quad \mathbf{x}, \mathbf{y} \in \mathbb{S}^d,$$

which becomes a metric if the kernel  $g$  (integrable on  $[0, 1]$ ) is positive but the proof of the corresponding invariance principle

$$\begin{aligned} & \frac{1}{N^2} \sum_{k, \ell=0}^{N-1} \rho(\mathbf{x}_k, \mathbf{x}_\ell) + 2 \int_{-1}^1 v(t) \int_{\mathbb{S}^d} \left| \sigma_d(C(\mathbf{x}; t)) - \frac{1}{N} \sum_{k=0}^{N-1} 1_{C(\mathbf{x}; t)}(\mathbf{x}_k) \right|^2 d\sigma_d(\mathbf{x}) \, dt \\ &= \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} \rho(\mathbf{x}, \mathbf{y}) \, d\sigma_d(\mathbf{x}) \, d\sigma_d(\mathbf{y}) \end{aligned} \quad (16)$$

making use of Haar integrals over the special orthogonal group  $\text{SO}(d+1)$  does not require it. Note that for  $g \equiv 1$  the function  $\rho(\mathbf{x}, \mathbf{y})$  is a constant multiple of the Euclidean distance. With some care one may even consider  $g(x) = 1/(1-x^2)$ .

It is well-known that a reproducing kernel  $K(\mathbf{x}, \mathbf{y})$  induces a distance (metric) by means of

$$d(\mathbf{x}, \mathbf{y}) = d_K(\mathbf{x}, \mathbf{y}) = \sqrt{K(\mathbf{x}, \mathbf{x}) - 2K(\mathbf{x}, \mathbf{y}) + K(\mathbf{y}, \mathbf{y})}.$$

For example, the reproducing kernel (7) yields

$$d_{K_{\mathcal{C}}}(\mathbf{x}, \mathbf{y}) = \sqrt{C_d} \sqrt{\|\mathbf{x} - \mathbf{y}\|}.$$

In general, for the symmetric weighted kernel  $K_{\mathcal{C}, v}(\mathbf{x}, \mathbf{y})$ , which does only depend on the inner product  $\langle \mathbf{x}, \mathbf{y} \rangle$ , it follows that ( $\mathbf{a} \in \mathbb{S}^d$  fixed)

$$K_{\mathcal{C}, v}(\mathbf{x}, \mathbf{y}) = \frac{1}{2} \left\{ [d_{\mathcal{C}, v}(\mathbf{x}, \mathbf{y})]^2 - K_{\mathcal{C}, v}(\mathbf{x}, \mathbf{x}) - K_{\mathcal{C}, v}(\mathbf{y}, \mathbf{y}) \right\} = \frac{1}{2} [d_{\mathcal{C}, v}(\mathbf{x}, \mathbf{y})]^2 - K_{\mathcal{C}, v}(\mathbf{a}, \mathbf{a}).$$

By Theorem 1 one arrives at

$$\begin{aligned} & \frac{1}{N^2} \sum_{k, \ell=0}^{N-1} [d_{\mathcal{C}, v}(\mathbf{x}_k, \mathbf{x}_\ell)]^2 + 2 \int_{-1}^1 v(t) \int_{\mathbb{S}^d} \left| \sigma_d(C(\mathbf{x}; t)) - \frac{1}{N} \sum_{k=0}^{N-1} 1_{C(\mathbf{x}; t)}(\mathbf{x}_k) \right|^2 d\sigma_d(\mathbf{x}) \, dt \\ &= \int_{\mathbb{S}^d} \int_{\mathbb{S}^d} [d_{\mathcal{C}, v}(\mathbf{x}, \mathbf{y})]^2 d\sigma_d(\mathbf{x}) d\sigma_d(\mathbf{y}), \end{aligned}$$

which should be compared with (16).

**Acknowledgement:** The first author is grateful to the School of Mathematics and Statistics at UNSW for their support.

## A Auxiliary results

The normalized surface area measure  $\sigma_d$  on  $\mathbb{S}^d$  admits the following decomposition

$$d\sigma_d(\mathbf{y}) = \frac{\omega_{d-1}}{\omega_d} (1-t^2)^{d/2-1} dt d\sigma_{d-1}(\mathbf{y}^*), \quad \mathbf{y} = (\sqrt{1-t^2} \mathbf{y}^*, t) \in \mathbb{S}^d, \quad (17)$$



where  $t \in [-1, 1]$ ,  $\mathbf{y}^* \in \mathbb{S}^{d-1}$  and  $\omega_d$  denotes the surface area of  $\mathbb{S}^d$  (cf Müller [11]). (By definition  $\langle \mathbf{y}, \mathbf{p} \rangle = t$ , where  $\mathbf{p}$  is the North Pole of  $\mathbb{S}^d$ .) Thus, by rotational symmetry, the integral of a zonal function  $f(\langle \mathbf{z}, \cdot \rangle)$ ,  $\mathbf{z} \in \mathbb{S}^d$  fixed, with respect to  $\sigma_d$  reduces to

$$\int_{\mathbb{S}^d} f(\langle \mathbf{z}, \mathbf{y} \rangle) d\sigma_d(\mathbf{y}) = \int_{\mathbb{S}^d} f(\langle \mathbf{p}, \mathbf{y} \rangle) d\sigma_d(\mathbf{y}) = \frac{\omega_{d-1}}{\omega_d} \int_{-1}^1 f(t) (1-t^2)^{d/2-1} dt.$$

*Proof.* [Proof of relations (6)] One gets

$$\begin{aligned} C_d &= \frac{1}{2} \int_{\mathbb{S}^d} |\langle \mathbf{p}, \mathbf{y} \rangle| d\sigma_d(\mathbf{y}) = \frac{1}{2} \frac{\omega_{d-1}}{\omega_d} \int_{-1}^1 |t| (1-t^2)^{d/2-1} dt = \frac{1}{2} \frac{\omega_{d-1}}{\omega_d} \int_0^1 (1-t^2)^{d/2-1} 2t dt \\ &= \frac{1}{d} \frac{\omega_{d-1}}{\omega_d} = \frac{1}{d} \frac{\Gamma((d+1)/2)}{\sqrt{\pi} \Gamma(d/2)} = \frac{\mathcal{H}_d(\mathbb{B}^d)}{\mathcal{H}_d(\mathbb{S}^d)} \sim \frac{1}{\sqrt{2\pi d}} \quad \text{as } d \rightarrow \infty. \end{aligned}$$

The second equality follows from

$$1 = \sigma_d(\mathbb{S}^d) = \frac{\omega_{d-1}}{\omega_d} \int_{-1}^1 (1-t^2)^{d/2-1} dt = \frac{\omega_{d-1}}{\omega_d} \int_0^1 v^{1/2-1} (1-v)^{d/2-1} dv = \frac{\omega_{d-1}}{\omega_d} B(1/2, d/2),$$

where  $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$  is the beta function. The third equality follows from the well-known formulas for the volume of the unit ball in  $\mathbb{R}^d$  and the surface area of  $\mathbb{S}^d$ . The asymptotics follows from the asymptotic expansion of a ratio of Gamma functions (cf. [1]).  $\square$

## B The weighted reproducing kernel

Next, we investigate the weighted reproducing kernel (13) in more detail. In particular, it will be shown that the kernel  $K_{\mathcal{E},v}(\mathbf{x}, \mathbf{y})$  is a function of the inner product  $\langle \mathbf{x}, \mathbf{y} \rangle$ .

On observing that  $\langle \mathbf{x}, \mathbf{z} \rangle \leq \langle \mathbf{y}, \mathbf{z} \rangle$  if and only if  $\langle \mathbf{y} - \mathbf{x}, \mathbf{z} \rangle \geq 0$  we may write for  $\mathbf{x} \neq \mathbf{y}$

$$K_{\mathcal{E},v}(\mathbf{x}, \mathbf{y}) = \mathcal{A}_{\mathcal{E},v}(\mathbf{x}, \mathbf{y}) + \mathcal{A}_{\mathcal{E},v}(\mathbf{y}, \mathbf{x}) - V(-1),$$

which immediately shows symmetry of the reproducing kernel, where

$$\mathcal{A}_{\mathcal{E},v}(\mathbf{x}, \mathbf{y}) = \int_{\mathbb{S}^d} V(\langle \mathbf{x}, \mathbf{z} \rangle) 1_{[0,1]}(\langle \frac{\mathbf{y} - \mathbf{x}}{\|\mathbf{y} - \mathbf{x}\|}, \mathbf{z} \rangle) d\sigma_d(\mathbf{z}). \quad (18)$$

By abuse of notation we set (note that  $\langle \mathbf{x}, \mathbf{y} \rangle = u$ )

$$\begin{aligned} \mathbf{z} &= t\mathbf{x} + \sqrt{1-t^2}\mathbf{z}^*, & -1 \leq t \leq 1, \mathbf{z}^* \in \mathbb{S}^{d-1}, \\ \mathbf{y} &= u\mathbf{x} + \sqrt{1-u^2}\mathbf{y}^*, & -1 \leq t \leq 1, \mathbf{y}^* \in \mathbb{S}^{d-1}. \end{aligned}$$

In this way  $\mathbf{x}$  will be the 'North Pole' in the decomposition (17) and we obtain

$$\mathcal{A}_{\mathcal{E},v}(\mathbf{x}, \mathbf{y}) = \frac{\omega_{d-1}}{\omega_d} \int_{-1}^1 V(t) \left\{ \int_{\mathbb{S}^{d-1}} 1_{[0,1]}(\langle \frac{\mathbf{y} - \mathbf{x}}{\|\mathbf{y} - \mathbf{x}\|}, \mathbf{z} \rangle) \sigma_{d-1}(\mathbf{z}^*) \right\} (1-t^2)^{d/2-1} dt.$$

The indicator function in the inner integral is a zonal function depending on  $w = \langle \mathbf{y}^*, \mathbf{z}^* \rangle$  only. Thus, we apply again (17) with  $\mathbf{y}^*$  as 'North Pole'. That is

$$\mathcal{A}_{\mathcal{E},v}(\mathbf{x}, \mathbf{y}) = \frac{\omega_{d-1}}{\omega_d} \int_{-1}^1 V(t) \left\{ \frac{\omega_{d-2}}{\omega_{d-1}} \int_{-1}^1 1_{[0,1]}(\langle \frac{\mathbf{y} - \mathbf{x}}{\|\mathbf{y} - \mathbf{x}\|}, \mathbf{z} \rangle) (1-w^2)^{(d-1)/2-1} dw \right\} (1-t^2)^{d/2-1} dt,$$

where the inner product evaluates as

$$\left\langle \frac{\mathbf{y} - \mathbf{x}}{\|\mathbf{y} - \mathbf{x}\|}, \mathbf{z} \right\rangle = \sqrt{1 - t^2} \sqrt{\frac{1 + \langle \mathbf{x}, \mathbf{y} \rangle}{2}} w - \sqrt{\frac{1 - \langle \mathbf{x}, \mathbf{y} \rangle}{2}} t, \quad \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{S}^d, \mathbf{x} \neq \mathbf{y}. \quad (19)$$

Proceeding similarly for  $\mathcal{A}_{\mathcal{C},v}(\mathbf{y}, \mathbf{x})$ , one sees that, indeed,  $\mathcal{A}_{\mathcal{C},v}(\mathbf{y}, \mathbf{x}) = \mathcal{A}_{\mathcal{C},v}(\mathbf{x}, \mathbf{y})$ . Furthermore,  $\mathcal{A}_{\mathcal{C},v}(\mathbf{x}, \mathbf{y})$  depends only on the inner product  $\langle \mathbf{x}, \mathbf{y} \rangle$  which in turn implies that the reproducing kernel  $K_{\mathcal{C},v}(\mathbf{x}, \mathbf{y})$  is a function of the inner product  $\langle \mathbf{x}, \mathbf{y} \rangle$ .

The right-hand side in (19) describes a line which stays strictly between the levels  $-1$  and  $1$  for  $v$  in  $[-1, 1]$  by the left-hand side in (19). Further analysis gives that the indicator functions is one (i) if  $t \leq -\sqrt{(1+u)/2}$  and  $-1 \leq v \leq 1$ , or, (ii) if  $-\sqrt{(1+u)/2} \leq t \leq \sqrt{(1+u)/2}$  and  $(t/\sqrt{1-t^2})\sqrt{(1-u)/(1+u)} \leq v \leq 1$ , and zero otherwise. This leads to

$$\begin{aligned} \mathcal{A}_{\mathcal{C},v}(\mathbf{x}, \mathbf{y}) &= \frac{\omega_{d-1}}{\omega_d} \int_{-1}^{-\sqrt{(1+u)/2}} V(t) (1 - t^2)^{d/2-1} dt \\ &\quad + \frac{\omega_{d-1}}{\omega_d} \int_{-\sqrt{(1+u)/2}}^{\sqrt{(1+u)/2}} V(t) I_{(1-x(t))/2}((d-1)/2, (d-1)/2) (1 - t^2)^{d/2-1} dt, \end{aligned}$$

where, when using  $u = \langle \mathbf{x}, \mathbf{y} \rangle = \cos \phi$  ( $0 < \phi < \pi$ ) and  $t = \cos \psi$ , one has

$$\sqrt{\frac{1+u}{2}} = \cos(\phi/2), \quad \sqrt{\frac{1-u}{1+u}} = \tan(\phi/2), \quad x(t) = \sqrt{\frac{1-u}{1+u}} \frac{t}{\sqrt{1-t^2}} = \frac{\cot \psi}{\cot(\phi/2)}.$$

The change of variable  $\xi = x(t)$  yields

$$\begin{aligned} \mathcal{A}_{\mathcal{C},v}(\mathbf{x}, \mathbf{y}) &= \frac{\omega_{d-1}}{\omega_d} \int_{-1}^{-\sqrt{(1+u)/2}} V(t) (1 - t^2)^{d/2-1} dt \\ &\quad + \frac{\omega_{d-1}}{\omega_d} \left( \frac{1-u}{1+u} \right)^{d/2} \int_{-1}^1 V\left( \frac{\xi}{\sqrt{\frac{1-u}{1+u} + \xi^2}} \right) I_{(1-\xi)/2}((d-1)/2, (d-1)/2) \frac{d\xi}{\left( \frac{1-u}{1+u} + \xi^2 \right)^{(d+1)/2}}, \end{aligned}$$

where we make use of the regularized incomplete beta function

$$I_z(a, b) = B_z(a, b) / B(a, b), \quad B_z(a, b) = \int_0^z t^{a-1} (1-t)^{b-1} dt, \quad a, b > 0.$$

We compute the following integral (using (18)):

$$\int_{\mathbb{S}^d} \mathcal{A}_{\mathcal{C},v}(\mathbf{x}, \mathbf{y}) d\sigma_d(\mathbf{y}) = \int_{\mathbb{S}^d} V(\langle \mathbf{x}, \mathbf{z} \rangle) \int_{\mathbb{S}^d} 1_{[0,1]} \left( \left\langle \frac{\mathbf{y} - \mathbf{x}}{\|\mathbf{y} - \mathbf{x}\|}, \mathbf{z} \right\rangle \right) d\sigma_d(\mathbf{y}) d\sigma_d(\mathbf{z}).$$

The inner integral is one if  $\mathbf{z}$  is in the half-sphere centered at  $-\mathbf{x}$  and zero otherwise. Hence,

$$\int_{\mathbb{S}^d} \mathcal{A}_{\mathcal{C},v}(\mathbf{x}, \mathbf{y}) d\sigma_d(\mathbf{y}) = \frac{\omega_{d-1}}{\omega_d} \int_0^1 V(-t) (1 - t^2)^{d/2-1} dt.$$

Since  $g'(t) = (\omega_{d-1}/\omega_d)(1 - t^2)^{d/2-1}$  for  $g(t) = (1/2) I_{t^2}(1/2, d/2)$ ,  $0 \leq t \leq 1$ , integration by parts gives

$$\int_{\mathbb{S}^d} \mathcal{A}_{\mathcal{C},v}(\mathbf{x}, \mathbf{y}) d\sigma_d(\mathbf{y}) = \frac{1}{2} V(-1) + \frac{1}{2} \int_0^1 v(-t) I_{t^2}(1/2, d/2) dt.$$

It follows that

$$\int_{\mathbb{S}^d} \int_{\mathbb{S}^d} K_{\mathcal{C},v}(\mathbf{x}, \mathbf{y}) d\sigma_d(\mathbf{x}) d\sigma_d(\mathbf{y}) = 2 \frac{\omega_{d-1}}{\omega_d} \int_0^1 V(-t) (1 - t^2)^{d/2-1} dt - V(-1) \quad (20a)$$

$$= \int_0^1 v(-t) I_{t^2}(1/2, d/2) dt. \quad (20b)$$

## References

- [1] Digital Library of Mathematical Functions. 2010-05-07. National Institute of Standards and Technology from <http://dlmf.nist.gov/>.
- [2] N. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68:337–404, 1950.
- [3] J. Beck. Sums of distances between points on a sphere—an application of the theory of irregularities of distribution to discrete geometry. *Mathematika*, 31(1):33–41, 1984.
- [4] G. Björck. Distributions of positive mass, which maximize a certain generalized energy integral. *Ark. Mat.*, 3:255–269, 1956.
- [5] J. S. Brauchart. Note on a generalized invariance principle and its relevance for cap discrepancy and energy. In *Modern developments in multivariate approximation*, volume 145 of *Internat. Ser. Numer. Math.*, pages 41–55. Birkhäuser, Basel, 2003.
- [6] J. S. Brauchart. Invariance principles for energy functionals on spheres. *Monatsh. Math.*, 141(2):101–117, 2004.
- [7] J. S. Brauchart and R. S. Womersley. Numerical integration over the unit sphere,  $\mathbb{L}_2$ -discrepancy and sum of distances. Manuscript, 25 pages.
- [8] H. Cohn and A. Kumar. Universally optimal distribution of points on spheres. *J. Amer. Math. Soc.*, 20(1):99–148, 2007.
- [9] J. Cui and W. Freeden. Equidistribution on the sphere. *SIAM J. Sci. Comput.*, 18(2):595–609, 1997.
- [10] D. P. Hardin and E. B. Saff. Discretizing manifolds via minimum energy points. *Notices Amer. Math. Soc.*, 51(10):1186–1194, 2004.
- [11] C. Müller. *Spherical harmonics*, volume 17 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1966.
- [12] E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Volume II: Standard information for functionals*, volume 12 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2010.
- [13] E. B. Saff and A. B. J. Kuijlaars. Distributing many points on a sphere. *Math. Intelligencer*, 19(1):5–11, 1997.
- [14] W. M. Schmidt. Irregularities of distribution. IV. *Invent. Math.*, 7:55–82, 1969.
- [15] I. H. Sloan and R. S. Womersley. Extremal systems of points and numerical integration on the sphere. *Adv. Comput. Math.*, 21(1-2):107–125, 2004.
- [16] K. B. Stolarsky. Sums of distances between points on a sphere. II. *Proc. Amer. Math. Soc.*, 41:575–582, 1973.
- [17] K. B. Stolarsky. Discrepancy and sums of distances between points of a metric space. In *The geometry of metric and linear spaces (Proc. Conf., Michigan State Univ., East Lansing, Mich., 1974)*, pages 44–56. Lecture Notes in Math., Vol. 490. Springer, Berlin, 1975.